



OY 1 ENKHNKCLQSCNSERDSTRNQCACHARNLKYKECECEGELPPRRPQHPER 55  
 DB 31 ENKHNKCLQSCNSERDSTRNQCACHARNLKYKECECEGELPPRRPQHPER 85

## RESULT 2

ID GLCX\_SOYBN STANDARD: PRT: 639 AA.  
 AC P11827;  
 DT 01-OCT-1989 (rel. 12, Created)  
 DT 01-OCT-1989 (rel. 12, Last sequence update)  
 DT 01-AUG-1992 (rel. 23, Last annotation update)  
 DE BETA-CONGLYCININ, ALPHA' CHAIN PRECURSOR.  
 GN CG-1.  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Fabales; Fabaceae; Papilionoideae; Glycine.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86250867; PubMed=3013879;  
 RA Doyle J.J., Schuller M.A., Godette W.D., Zenger V., Beachy R.N.,  
 RA Slightom J.L.;  
 RT "The glycosylated seed storage proteins of Glycine max and Phaseolus  
 vulgaris. Structural homologues of genes and proteins.";  
 RL J. Biol. Chem. 261:9228-9238(1986).  
 RN [2]  
 RP SEQUENCE OF 340-639 FROM N.A.  
 RX MEDLINE=83143289; PubMed=6897678;  
 RA Schuller M.A., Ladin B.F., Pollaco J.C., Freyer G., Beachy R.N.;  
 RT "Structural sequences are conserved in the genes coding for the  
 alpha, alpha' and beta-subunits of the soybean 7S seed storage  
 protein.";  
 RL Nucleic Acids Res. 10:8245-8261(1982).  
 CC -1- FUNCTION: SEED STORAGE PROTEIN. ACCUMULATES DURING SEED  
 DEVELOPMENT AND IS HYDROLYZED AFTER GERMINATION TO PROVIDE A  
 CARBON AND NITROGEN SOURCE FOR THE DEVELOPING SEEDLING.  
 CC -1- SUBUNIT: THE ALPHA', ALPHA'', AND BETA-SUBUNITS ASSOCIATE IN  
 VARIOUS COMBINATIONS TO FORM TRIMERIC PROTEINS.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC MEMBRANE-BOUND VACUOLAR PROTEIN  
 BODIES.  
 CC -1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,  
 CONVICLIN, CONGLYCININ, ETC.).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M13759; AAB01374.1; -  
 DR EMBL: J01290; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: B24810; B24810.  
 DR HSSP: P02853; 2PHL.  
 DR INTERPRO: IPR001113; -  
 DR PFAM: PF00546; Seedstore\_7s; 2.  
 DR SEED storage protein; Signal; Glycoprotein; Multigene family.  
 KW SIGNAL  
 FT PROPEP 1 22  
 FT CHAIN 23 62  
 FT CARBOHYD 63 639 BETA-CONGLYCININ, ALPHA' CHAIN.  
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 551 551 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 543 543 P -> L (IN REF. 2).  
 FT CONFLICT 549 549 M -> V (IN REF. 2).  
 FT CONFLICT 608 608 S -> T (IN REF. 2).  
 SQ SEQUENCE 639 AA: 74325 MW: 469BF24C79651E3F CRC64:

Query Match 82.4%; Score 259.5; DB 1: Length 639;  
 Best Local Similarity 83.9%; Pred. No. 1.3e-20;  
 Matches 47; Conservative 5; Mismatches 1; Indels 3; Gaps 2;

OY 1 ENKHNKCLQSCNSERDSTRNQCACHARNLKY-KECECEGELPPRRPQHPER 55  
 DB 31 QNPSHNKCLQSCNSERDSTRNQCACHARNLKYEEBCECEGEL--PPRRPQHPER 84

## RESULT 3

ID AH12\_ARAHY STANDARD: PRT: 614 AA.  
 AC P43237;  
 DT 01-NOV-1995 (rel. 32, Created)  
 DT 01-NOV-1995 (rel. 32, Last sequence update)  
 DT 01-NOV-1997 (rel. 35, Last annotation update)  
 DE ALLERGEN ARA H 1, CLONE P17 (ARA H 1).  
 OS Arachis hypogaea (Peanut).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Fabales; Fabaceae; Papilionoideae; Arachis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. FLORUNNER;  
 RX MEDLINE=96013631; PubMed=7560062;  
 RA Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;  
 RT "Recombinant peanut allergen Ara h I expression and IGE binding in  
 patients with peanut hypersensitivity.";  
 RL J. Clin. Invest. 96:1715-1721(1995).  
 CC -1- SIMILARITY: TO 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,  
 CONVICLIN, CONGLYCININ, ETC.).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U38853; AAA60336.1; -  
 DR HSSP: P50477; ICAM.  
 DR INTERPRO: IPR001113; -  
 DR PFAM: PF00546; Seedstore\_7s; 1.  
 KW Allergen.  
 SQ SEQUENCE 614 AA: 70283 MW: 1DDACF217EBC5F31 CRC64:

Query Match 28.3%; Score 89; DB 1: Length 614;  
 Best Local Similarity 44.1%; Pred. No. 0.017;  
 Matches 15; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

OY 1 ENKHNKCLQSCNSERDSTRNQCACHARNLKYVE 34  
 DB 33 ENPCARCLQSCNSERDSTRNQCACHARNLKYVE 66

## RESULT 4

ID AH12\_ARAHY STANDARD: PRT: 626 AA.  
 AC P43238;  
 DT 01-NOV-1995 (rel. 32, Created)  
 DT 01-NOV-1995 (rel. 32, Last sequence update)  
 DT 01-NOV-1997 (rel. 35, Last annotation update)  
 DE ALLERGEN ARA H 1, CLONE P41B (ARA H 1).  
 OS Arachis hypogaea (Peanut).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Fabales; Fabaceae; Papilionoideae; Arachis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. FLORUNNER;  
 RX MEDLINE=96013631; PubMed=7560062;  
 RA Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;  
 RT "Recombinant peanut allergen Ara h I expression and IGE binding in  
 patients with peanut hypersensitivity.";

RL J. Clin. Invest. 96:1715-1721(1995).  
 CC - SIMILARITY: TO 7S SEED STORAGE PROTEINS (PHASCOLIN, VICILIN,  
 CC CONVICILIN, CONGLYCININ, ETC.).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/sib.ch>).  
 CC or send an email to [license@sib.slb.ch](mailto:license@sib.slb.ch)).  
 CC -----  
 DR EMBL: L34402; AAB00861.1; -.  
 DR HSSP: P50477; ICAM.  
 DR INTERPRO: IPR001113. -.  
 DR PFAM: PF00546; Seedstore\_7s; 1.  
 KW Allergen.  
 SO SEQUENCE. 626 AA; 71345 MW; 1A6BBBE41490D0E3 CRC64;

|                       |        |                  |       |                |
|-----------------------|--------|------------------|-------|----------------|
| Query Match           | 28.3%; | Score 89;        | DB 1; | Length 626;    |
| Best Local Similarity | 44.1%; | Pred. NO. 0.017; |       |                |
| Matches               | 15;    | Conservative     | 6;    | Mismatches 13; |
|                       |        |                  |       | Indels 0;      |
|                       |        |                  |       | Gaps 0;        |

```
QY      1 ENPRHNKCLÖSCNSERDSYRNÖACHARNLLKVE 34
      ||| :||| | | : :|| :| : :
Db      35 ENPCAÖRCLÖSCQÖEPDDLKÖKACESRCTKLEYD 68
```

|            |            |           |
|------------|------------|-----------|
| RESULT     | 5          |           |
| TONB_PASHA |            |           |
| ID         | TONB_PASHA | STANDARD; |
| 7700       |            | PRT;      |
|            |            | 246 AA    |

|    |   |
|----|---|
| DT | 15-DEC-1998 (Rel. 37, Created)                |
| DT | 15-DEC-1998 (Rel. 37, last sequence update)   |
| DT | 30-MAY-2000 (Rel. 39, last annotation update) |

05 *Pasteurella haemolytica*.  
0C Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
0C *Pasteurella*.

RC STRAIN=SEROTYPE A1 / ATCC 43270;  
RA Graham M.R., Lo R.Y.C.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

CC THE PERIPLASMIC SPACE OF SPECIFIC SUBSTRATES. IT COULD ACT TO  
CC TRANSDUCE ENERGY FROM THE CYTOPLASMIC MEMBRANE TO SPECIFIC ENERGY-  
CC REQUIRING PROCESSES IN THE OUTER MEMBRANE, RESULTING IN THE  
CC DEFLECTION AND PROJECTION OF TONGUES TOWARD BY SURFACE CHARGES

CC -1- SUBCELLULAR LOCATION: ANCHORED TO THE CYTOPLASMIC  
CC MEMBRANE VIA ITS N-TERMINAL SIGNAL-LIKE SEQUENCE, SPANS THE  
CC PERIPLASM (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE TOXN FAMILY

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on  
CC use by non-profit institutions as long as the source is acknowledged.

-----  
 CC entitisrequires a license agreement (see [http://www.isb-sib.ch/announcement/](http://www.isb-sib.ch/announcement)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC  
 DR EMRI: M62565: AAR09530 1. -  
 -----

|    |          |         |           |                          |
|----|----------|---------|-----------|--------------------------|
| FT | TRANSMEM | 8       | 28        | ANCHOR (POTENTIAL).      |
| FT | DOMAIN   | 29      | 246       | PERIPLASMIC (POTENTIAL). |
| SQ | SEQUENCE | 246 AA; | 27785 MW; | C9582F619FCBA5B5 CRC64;  |

|                       |       |                 |            |                                  |
|-----------------------|-------|-----------------|------------|----------------------------------|
| Query Match           | 21.9% | Score 69        | DB 1:      | Length 246                       |
| Best Local Similarity | 32.1% | Pred. No. 0.99% |            |                                  |
| Matches               | 17    | Conservative    | 11         | Mismatches 17; Indels 8; Gaps 2  |
| QY                    | 2     | NPKNKLCQSCNS    | EDSTRNQAC  | CHARNILKYEKECEGEGTLP--RPRRPOH 52 |
|                       |       |                 |            |                                  |
| DB                    | 100   | NPKKKPKPEKKRR   | HHKHHKHH-- | -----KKRRLELOQEPKPKPKPERPH 146   |

|            |            |           |
|------------|------------|-----------|
| RESULT     | 6          |           |
| VCLB_GOSHI |            |           |
| ID         | VCLB_GOSHI | STANDARD; |
|            |            | PRT;      |
|            |            | 588 AA.   |

DT 01-MAR-1989 (Rel. 10, last sequence update)  
DT 15-JUL-1999 (Rel. 38, last annotation update)  
DE VICILIN C72 PRECURSOR (ALPHA-GLOBULIN B).  
Occasional literature references

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II,  
OC Malvales; Malvaceae; Gossypium.  
RN [1]  
RD SOURCE: NCBI

RT "Developmental biochemistry of cottonseed embryogenesis and  
germination. XVIII. cDNA and amino acid sequences of the members of  
the storage protein families.";  
Plant Mol Biol 7:475-480(1986)

CC -1- SUBCELLULAR LOCATION: CYTOSOL  
CC BODIES.  
CC  
CC -1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONVICTILIN, CONGLUTININ ETC.)  
CC  
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC at the European Bioinformatics Institute. There are no restrictions on  
CC use by non-profit institutions as long as its content is in no way  
CC altered.

EMBL: M16891: AAA33071.1: -  
 CC enucleates requires a license agreement (see [http://www.isd-sib.ch/announcement/or\\_send\\_an\\_email\\_to\\_license@isd-sib.ch](http://www.isd-sib.ch/announcement/or_send_an_email_to_license@isd-sib.ch)).  
 CC  
 CC  
 CC  
 DR

```
DR HSSP; P50477; ICAM.  
DR INTERPRO; IPR001113; -.  
DR PFAM; PF00546; Seedstore_7s; 1.  
KW Seed storage protein; Signal.
```

| FI | CHAIN    | 20      | 588       | VICLIN C/2.             |
|----|----------|---------|-----------|-------------------------|
| SQ | SEQUENCE | 588 AA; | 69729 MW; | 63E699B29AB8ADEB CRC64, |

|    |   |   |             |                |
|----|---|---|-------------|----------------|
| 02 | 1 | END- <del>KHKKCLO</del> CNSERDYSYRQACHARCINLTKVEKE----- | ECEEGETPRPR | 49             |
|    |   | Best Local Similarity                                   | 26.0%       | Ident. No. 11; |
|    |   | Matches   | 19;         | Mismatches     |
|    |   |   | 19;         | Indels         |
|    |   |   |             | Gaps           |
|    |   |   |             | 5;             |

01 LDPGNLDECECNDVENVNFY -COON- UNIT EXEEOOOSORFVECOO ---HCHQQ 134  
DZ  
QY 50 PÖPPER 55  
| | |

Db 135 EQRPER 140

|            |                                |
|------------|--------------------------------|
| INM3_CAEEL |                                |
| ID         | YNX3_CAEEL                     |
| AC         | STANDARD;                      |
|            | P34576;                        |
| DT         | 01-FEB-1994 (Rel. 28, Created) |
|            | PRT; 3051 AA                   |

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, last sequence update)  
DE 01-NOV-1997 (Rel. 35, last annotation update)  
GN HYPOTHETICAL PROTEIN T20G5.3 IN CHROMOSOME III (FRAGMENT).  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderae; Caenorhabditis.  
RN [1]  
RC SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Berks M., Smith A.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 45 EGF-LIKE DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z30423; CAA83007.1; -.  
CC PIR: S42373; S42373.  
DR HSSP: P35555; IEMO.  
DR WORMPEP: T20G5.3; CE00478.  
DR INTERPRO: IPR000082; -.  
DR INTERPRO: IPR000152; -.  
DR INTERPRO: IPR000561; -.  
DR INTERPRO: IPR001881; -.  
DR INTERPRO: IPR002035; -.  
DR PFAM: PF00008; EGF\_31.  
DR PFAM: PF001390; SEA; 2.  
DR PFAM: PF00092; vwa; 1.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 28.  
DR PROSITE: PS01186; EGF\_2; 5.  
DR PROSITE: PS01187; EGF\_CA; 1.  
KW Hypothetical protein; EGF-like domain; Repeat; Transmembrane.  
FT NON\_TER 1  
FT DOMAIN <1 2701 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2702 2722 POTENTIAL.  
FT DOMAIN 2723 3051 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN <1 44 EGF-LIKE 1 (PARTIAL).  
FT DOMAIN 46 94 EGF-LIKE 2.  
FT DOMAIN 100 144 EGF-LIKE 3.  
FT DOMAIN 145 192 EGF-LIKE 4.  
FT DOMAIN 194 245 EGF-LIKE 5.  
FT DOMAIN 247 296 EGF-LIKE 6.  
FT DOMAIN 313 354 EGF-LIKE 7.  
FT DOMAIN 355 402 EGF-LIKE 8.  
FT DOMAIN 404 452 EGF-LIKE 9.  
FT DOMAIN 454 503 EGF-LIKE 10.  
FT DOMAIN 514 699 VWA.  
FT DOMAIN 705 750 EGF-LIKE 11.  
FT DOMAIN 750 794 EGF-LIKE 12.  
FT DOMAIN 805 846 EGF-LIKE 13.  
FT DOMAIN 847 892 EGF-LIKE 14.  
FT DOMAIN 892 940 EGF-LIKE 15.  
FT DOMAIN 942 990 EGF-LIKE 16.  
FT DOMAIN 992 1039 EGF-LIKE 17.  
FT DOMAIN 1043 1091 EGF-LIKE 18.  
FT DOMAIN 1093 1144 EGF-LIKE 19.  
FT DOMAIN 1146 1195 EGF-LIKE 20.  
FT DOMAIN 1197 1245 EGF-LIKE 21.  
FT DOMAIN 1247 1295 EGF-LIKE 22.  
FT DOMAIN 1298 1346 EGF-LIKE 23.  
FT DOMAIN 1352 1396 EGF-LIKE 24.  
FT DOMAIN 1397 1444 EGF-LIKE 25.  
FT DOMAIN 1446 1492 EGF-LIKE 26.  
FT DOMAIN 1494 1542 EGF-LIKE 27.  
FT DOMAIN 1544 1592 EGF-LIKE 28.  
FT DOMAIN 1594 1642 EGF-LIKE 29.

FT DOMAIN 1644 1692 EGF-LIKE 30.  
FT DOMAIN 1693 1739 EGF-LIKE 31.  
FT DOMAIN 1740 1788 EGF-LIKE 32.  
FT DOMAIN 1797 1847 EGF-LIKE 33.  
FT DOMAIN 1849 1900 EGF-LIKE 34.  
FT DOMAIN 1902 1950 EGF-LIKE 35.  
FT DOMAIN 1952 1998 EGF-LIKE 36.  
FT DOMAIN 2000 2047 EGF-LIKE 37.  
FT DOMAIN 2047 2095 EGF-LIKE 38.  
FT DOMAIN 2117 2156 EGF-LIKE 39.  
FT DOMAIN 2155 2284 SEA 1.  
FT DOMAIN 2293 2332 EGF-LIKE 40.  
FT DOMAIN 2331 2457 SEA 2.  
FT DOMAIN 2460 2504 EGF-LIKE 41.  
FT DOMAIN 2508 2556 EGF-LIKE 42.  
FT DOMAIN 2556 2608 EGF-LIKE 43.  
FT DOMAIN 2612 2657 EGF-LIKE 44, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 2657 2693 EGF-LIKE 45.  
FT DISULFID 2  
FT DISULFID 7  
FT DISULFID 24  
FT DISULFID 50  
FT DISULFID 57  
FT DISULFID 74  
FT DISULFID 104  
FT DISULFID 112  
FT DISULFID 131  
FT DISULFID 149  
FT DISULFID 157  
FT DISULFID 174  
FT DISULFID 198  
FT DISULFID 208  
FT DISULFID 225  
FT DISULFID 251  
FT DISULFID 259  
FT DISULFID 276  
FT DISULFID 317  
FT DISULFID 324  
FT DISULFID 341  
FT DISULFID 359  
FT DISULFID 365  
FT DISULFID 382  
FT DISULFID 408  
FT DISULFID 415  
FT DISULFID 432  
FT DISULFID 458  
FT DISULFID 466  
FT DISULFID 483  
FT DISULFID 709  
FT DISULFID 717  
FT DISULFID 736  
FT DISULFID 749  
FT DISULFID 754  
FT DISULFID 762  
FT DISULFID 780  
FT DISULFID 809  
FT DISULFID 816  
FT DISULFID 833  
FT DISULFID 851  
FT DISULFID 859  
FT DISULFID 876  
FT DISULFID 878  
FT DISULFID 891  
FT DISULFID 896  
FT DISULFID 903  
FT DISULFID 920  
FT DISULFID 939  
FT DISULFID 946  
FT DISULFID 953  
FT DISULFID 970  
FT DISULFID 989  
FT DISULFID 996  
FT DISULFID 1004  
FT DISULFID 1021  
FT DISULFID 1047  
FT DISULFID 1054  
FT DISULFID 1072  
FT DISULFID 1097  
FT DISULFID 1113

```

FT DISULFID 1105 1122 BY SIMILARITY.
FT DISULFID 1124 1143 BY SIMILARITY.
FT DISULFID 1150 1164 BY SIMILARITY.
FT DISULFID 1157 1173 BY SIMILARITY.
FT DISULFID 1175 1194 BY SIMILARITY.
FT DISULFID 1201 1214 BY SIMILARITY.
FT DISULFID 1208 1223 BY SIMILARITY.
FT DISULFID 1225 1244 BY SIMILARITY.
FT DISULFID 1251 1264 BY SIMILARITY.
FT DISULFID 1258 1273 BY SIMILARITY.
FT DISULFID 1275 1294 BY SIMILARITY.
FT DISULFID 1302 1315 BY SIMILARITY.
FT DISULFID 1309 1324 BY SIMILARITY.
FT DISULFID 1326 1345 BY SIMILARITY.
FT DISULFID 1356 1372 BY SIMILARITY.
FT DISULFID 1364 1381 BY SIMILARITY.
FT DISULFID 1383 1395 BY SIMILARITY.
FT DISULFID 1401 1415 BY SIMILARITY.
FT DISULFID 1409 1424 BY SIMILARITY.
FT DISULFID 1426 1443 BY SIMILARITY.
FT DISULFID 1450 1464 BY SIMILARITY.
FT DISULFID 1458 1473 BY SIMILARITY.
FT DISULFID 1475 1491 BY SIMILARITY.
FT DISULFID 1498 1512 BY SIMILARITY.
FT DISULFID 1506 1521 BY SIMILARITY.
FT DISULFID 1523 1541 BY SIMILARITY.
FT DISULFID 1548 1562 BY SIMILARITY.
FT DISULFID 1556 1571 BY SIMILARITY.
FT DISULFID 1573 1591 BY SIMILARITY.
FT DISULFID 1598 1611 BY SIMILARITY.
FT DISULFID 1605 1620 BY SIMILARITY.
FT DISULFID 1622 1641 BY SIMILARITY.
FT DISULFID 1648 1661 BY SIMILARITY.
FT DISULFID 1655 1670 BY SIMILARITY.
FT DISULFID 1672 1691 BY SIMILARITY.
FT DISULFID 1697 1709 BY SIMILARITY.
FT DISULFID 1703 1719 BY SIMILARITY.
FT DISULFID 1721 1738 BY SIMILARITY.
FT DISULFID 1744 1758 BY SIMILARITY.
FT DISULFID 1752 1767 BY SIMILARITY.
FT DISULFID 1769 1787 BY SIMILARITY.
FT DISULFID 1801 1815 BY SIMILARITY.
FT DISULFID 1809 1825 BY SIMILARITY.
FT DISULFID 1827 1846 BY SIMILARITY.
FT DISULFID 1853 1867 BY SIMILARITY.
FT DISULFID 1861 1878 BY SIMILARITY.
FT DISULFID 1880 1899 BY SIMILARITY.
FT DISULFID 1906 1920 BY SIMILARITY.
FT DISULFID 1914 1929 BY SIMILARITY.
FT DISULFID 1931 1949 BY SIMILARITY.
FT DISULFID 1956 1970 BY SIMILARITY.

```

Query Match 19.8%; Score 62.5; DB 1; Length 3051;  
 Best Local Similarity 25.9%; Pred. No. 47;  
 Matches 15; Conservative 10; Mismatches 24; Indels 9; Gaps 3;

```

OY 2 NPKINKLOS--CNSERDYNQACHARNLTKVEKECEGEI---PPRPPRPP 53
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 1746 DPRINTCSRNAICDPRGR--CECKRGFMRSPPSSQKRGVCEPPPPPPPP 1800

RESULT 8
EAR2_MOUSE
ID PA3136: 061504: STANDARD; PRT; 390 AA.
AC PA3136: 061504:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ORPHAN NUCLEAR RECEPTOR EAR-2 (V-ERBA RELATED PROTEIN EAR-2).
GN NR2P6 OR ERBA2 OR EAR2 OR EAR-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95034311; PubMed=7947324;
RA Jonk L.J.C., de Jonge M.J., Pals C.E.G.M., Missink S.,
RA Verwaart J.M.A., Schoorlemmer J., Kruijer W.,
RT "Cloning and expression during development of three murine members of
RT the COUP family of nuclear orphan receptors";
RL Mech. Dev. 47:81-97(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PITUITARY;
RX MEDLINE=94252589; PubMed=8194772;
RA Barnhart K.M., Mellon P.L.;
RT "The sequence of a murine cDNA encoding Ear-2, a nuclear orphan
RT receptor.";
RL Gene 142:313-314(1994).
CC -|- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -|- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NR2 SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
-----
DR EMBL; X76554; CA54097.1; -.
DR EMBL; L25674; AAA37532.1; -.
DR HSSP; P19793; ZNL.
DR MGD; MGI:99530; ERBA2.
DR INTERPRO; IPR000536; -.
DR INTERPRO; IPR001628; -.
DR INTERPRO; IPR001723; -.
DR INTERPRO; IPR003068; -.
DR PFAM; PF00104; hormone_rec; 1.
DR PFAM; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROLDFFINGER.
DR PRINTS; PR00398; STROHOMONER.
DR PRINTS; PR01282; COUPRECTOR.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger.
FT DNA_BIND 57 122. C4-TYPE ZINC FINGERS (TWO).
FT ZN_FING 57 77. C4-TYPE.
FT ZN_FING 93 117. C4-TYPE.
FT CONFLICT 10 10. D -> G (IN REF. 2).
FT CONFLICT 152 152. P -> H (IN REF. 2).
FT CONFLICT 164 165. OL -> HV (IN REF. 2).
FT CONFLICT 196 198. LPL -> AV (IN REF. 2).
FT CONFLICT 239 241. LPL -> VAV (IN REF. 2).
FT CONFLICT 357 357. K -> Q (IN REF. 2).
SQ SEQUENCE 390 AA; 41982 MM; 4C271F439CAF157F CRC64;

```

Query Match 19.7%; Score 62; DB 1; Length 390;  
 Best Local Similarity 33.3%; Pred. No. 8.3;  
 Matches 16; Conservative 6; Mismatches 18; Indels 8; Gaps 2;

```

OY 11 SCNSERD-----SYRNQACHAR--CNILTKVEKECEGEIPIPPRPP 50
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 92 TCRSNRDCQIDQHNHRNCOYCRLLKCRFVGMRKKAAYGRGRIPIALRPP 139

RESULT 9
EAR2_RAT
ID EAR2_RAT STANDARD; PRT; 390 AA.
AC 009017:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ORPHAN NUCLEAR RECEPTOR EAR-2 (V-ERBA RELATED PROTEIN EAR-2)

```

```

DE (OVALBUMIN UPSTREAM PROMOTER GAMMA NUCLEAR RECEPTOR RCOUPG).
GN NR2F6 OR ERBA12 OR EAR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PIUITARY;
RA Boutin J.M., Ronsin B., Devost D., Lipkin S.M., Rosenfeld M.G.,
RA Morel G.;
RT "Cloning and expression of COUP and COUP in the rat anterior
RT pituitary.";
RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NR2 SUPERFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF003926; AAB61296.1; -.
DR HSSP; P19793; 2NL.
DR INTERPRO: IPR000536; -.
DR INTERPRO: IPR001628; -.
DR INTERPRO: IPR001723; -.
DR PFAM; PF00104; hormone_rec; 1.
DR PFAM; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00398; STROHORMONER.
DR PRINTS; PR01282; COUPNFACTOR.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KM Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KM Zinc-finger.
FT DNA_BIND 57 122 C4-TYPE ZINC FINGERS (TWO).
FT ZN_FING 57 77 C4-TYPE.
FT ZN_FING 93 117 C4-TYPE.
SQ SEQUENCE 390 AA: 41729 MW: B08FAC93F76A7783 CRC64;
Query Match 19.0%; Score 60; DB 1; Length 390;
Best Local Similarity 33.3%; Pred. No. 14;
Matches 16; Conservative 6; Mismatches 18; Indels 8; Gaps 2;
QY 11 SCNSEPD-----SYRNOACHR---CNLKVKEKECECEGIPRRPRP 50
DB 92 TCRSNRDCCQIDQHHNRCCOYCRLLKCFRVMRKEAVOPGP1PHALPGP 139
RESULT 10
CCAD_RAT STANDARD: PRT: 2203 AA.
AC P27732; Q63491; Q63492; Q62691; Q01542; Q09022; Q09023; Q09024;
AC Q62815;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE VOLTAGE-DEPENDENT L-TYPE CALCIUM CHANNEL ALPHA-1D SUBUNIT (CALCIUM
DE CHANNEL, L-TYPE, ALPHA-1 POLYPEPTIDE, ISOFORM 2) (RAT BRAIN CLASS D)
DE (RBD).
GN CACNA1D OR CACNA1A2 OR CCH1A2 OR CACH3 OR CACNA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS CACNA AND CACNA-B).
RC TISSUE=INSULINOMA.
RX MEDLINE=95280950; PubMed=7760845;
RA Ihara Y., Yamada Y., Fujii Y., Gonoi T., Yano H., Yasuda K.,

```

```

RA Inagaki N., Saino Y., Saino S.;
RT "Molecular diversity and functional characterization of voltage-
RT dependent calcium channels (CACNA) expressed in pancreatic beta-
RT cells.";
RL Mol. Endocrinol. 9:121-130(1995).
RN [2]
RP SEQUENCE FROM N.A. (TRUNCATED ISOFORM AND CLONES RB9/11/34/48).
RC TISSUE=BRAIN;
RX MEDLINE=91299338; PubMed=1648940;
RA Hul A., Ellinor P.T., Krizanov O., Wang J.-J., Diebold R.J.,
RA Schwartz A.;
RT "Molecular cloning of multiple subtypes of a novel rat brain isoform
RT of the alpha-1 subunit of the voltage-dependent calcium channel.";
RL Neuron 7:35-44(1991).
RN [3]
RP SEQUENCE OF 1-125 FROM N.A.
RX MEDLINE=96040125; PubMed=7553731;
RA Kamp T.J., Mitas M., Fields K.L., Asch S., Chin H., Marban E.,
RA Nirenberg M.;
RT "Transcriptional regulation of the neuronal L-type calcium channel
RT alpha 1D subunit gene.";
RL Cell. Mol. Neurobiol. 15:307-326(1995).
RN [4]
RP SEQUENCE OF 1100-1410 FROM N.A. (CLONES RKC5 AND RKC6).
RC TISSUE=KIDNEY;
RX MEDLINE=93066265; PubMed=1279681;
RA Yu A.S.L., Hebert S.C., Brenner B.M., Lytton J.;
RT "Molecular characterization and nephron distribution of a family of
RT transcripts encoding the pore-forming subunit of Ca2+ channels in the
RT kidney.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10494-10498(1992).
RN [5]
RP SEQUENCE OF 1218-1498 FROM N.A. (ISOFORM ROB3).
RC TISSUE=OSTEOSARCOMA;
RX MEDLINE=96074617; PubMed=7479909;
RA Barry E.L.R., Gesek P.A., Froehner S.C., Friedman P.A.;
RT "Multiple calcium channel transcripts in rat osteosarcoma cells:
RT selective activation of alpha 1D isoform by parathyroid hormone.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10914-10918(1995).
RN [6]
RP SEQUENCE OF 1200-1493 FROM N.A. (RH1; RH2; CACNA-B & DELTA-1VS3/54).
RC TISSUE=HEPATOMA;
RX MEDLINE=97376179; PubMed=9232351;
RA Breton H.M., Harland M.L., Froscio M., Pelronjic T.,
RA Barritt G.J.;
RT "Novel variants of voltage-operated calcium channel alpha-1 subunit
RT transcripts in a rat liver-derived cell line: deletion in the IVS4
RT voltage sensing region.";
RL Cell Calcium 22:35-52(1997).
RN [7]
RP SEQUENCE OF 1307-1479 FROM N.A. (CLONE RBD-55).
RX MEDLINE=90239020; PubMed=1692134;
RT Snutch T.P., Leonard J.P., Gilbert M.M., Lester H.A., Davidson N.;
RT "Rat brain expresses a heterogeneous family of calcium channels.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3391-3395(1990).
CC -1- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D
CC GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)
CC CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
CC GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),
CC PHENYLALKYLAMINES, BENZOTHIAZEPINES, AND BY OMEGA-AGATOXIN-111A
CC (OMEGA-AGA-111A). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-
CC GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).
CC -1- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.

```

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- ALTERNATIVE PRODUCTS: MULTIPLE ISOFORMS ARE CREATED BY ALTERNATIVE  
 CC SPLICING EVENTS, WHICH SEEM TO OCCUR IN A TISSUE-SPECIFIC MANNER.  
 CC THE SEQUENCE SHOWN HERE IS THAT OF CACNA4. THE REGION SEQUENCED IN  
 CC RO3 AND RK5 IS IDENTICAL TO CACNA.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, PANCREATIC ISLETS AND B-  
 CC LYMPHOCYTES.  
 CC -1- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE  
 CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S4, S5, S6) AND ONE  
 CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS  
 CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A  
 CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.  
 CC -1- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>  
 CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
 CC -----  
 DR EMBL; D38101; BAA07282.1; -;  
 DR EMBL; D38102; BAA07283.1; -;  
 DR EMBL; M57682; AAA42015.1; -;  
 DR EMBL; U14005; AAB60515.1; -;  
 DR EMBL; M99221; AAA40895.1; -;  
 DR EMBL; U31722; AAA89156.1; -;  
 DR EMBL; U49126; AAB61634.1; -;  
 DR EMBL; U49127; AAB61635.1; -;  
 DR EMBL; U49128; AAB61636.1; -;  
 DR INTERPRO: IPR000636; -;  
 DR INTERPRO: IPR002077; -;  
 DR INTERPRO: IPR003091; -;  
 DR PIRAM: PF00520; Ion\_trans; 4.  
 DR PRINTS: PR00167; CACHANNEL.  
 DR PRINTS: PR00169; KCHANNEL.  
 KM Ionic channel: Transmembrane: Ion transport: Voltage-gated channel;  
 KW Calcium channel; Glycoprotein; Repeat; Multigene family;  
 KW Calcium-binding; Phosphorylation; Alternative splicing.  
 FT REPEAT 112 408  
 FT REPEAT 528 774  
 FT REPEAT 892 1174  
 FT REPEAT 1211 1486  
 FT DOMAIN 1 126  
 FT TRANSSEM 127 145  
 FT DOMAIN 146 163  
 FT TRANSSEM 164 183  
 FT DOMAIN 184 195  
 FT TRANSSEM 196 214  
 FT DOMAIN 215 235  
 FT TRANSSEM 236 254  
 FT DOMAIN 255 273  
 FT TRANSSEM 274 293  
 FT DOMAIN 294 381  
 FT TRANSSEM 382 406  
 FT DOMAIN 407 582  
 FT TRANSSEM 583 602  
 FT DOMAIN 603 617  
 FT TRANSSEM 618 636  
 FT DOMAIN 637 644  
 FT TRANSSEM 645 663  
 FT DOMAIN 664 673  
 FT TRANSSEM 674 692  
 FT DOMAIN 693 711  
 FT TRANSSEM 712 732  
 FT DOMAIN 733 786  
 FT TRANSSEM 787 811  
 FT DOMAIN 812 945  
 FT TRANSSEM 946 964  
 FT DOMAIN 965 980  
 FT TRANSSEM 981 1000

FT DOMAIN 1001 1012 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 1013 1031 S3 OF REPEAT III (POTENTIAL).  
 FT DOMAIN 1032 1037 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 1038 1057 S4 OF REPEAT III (POTENTIAL).  
 FT DOMAIN 1058 1076 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 1077 1096 S5 OF REPEAT III (POTENTIAL).  
 FT DOMAIN 1097 1186 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 1187 1207 S6 OF REPEAT III (POTENTIAL).  
 FT DOMAIN 1208 1264 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 1265 1283 S1 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1284 1298 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 1299 1318 S2 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1319 1325 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 1326 1347 S3 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1348 1357 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 1358 1377 S4 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1378 1396 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 1397 1416 S5 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1417 1483 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 1484 1508 S6 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1509 2203 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 1 7 POLY-MET.  
 FT DOMAIN 712 718 POLY-LEU.  
 FT DOMAIN 886 897 BINDING TO THE BETA SUBUNIT (BY  
 FT DOMAIN 429 446 SIMILARITY).  
 FT SITE 364 364 CALCIUM ION SELECTIVITY AND PERMEABILITY  
 FT SITE 763 763 (BY SIMILARITY).  
 FT SITE 1160 1160 CALCIUM ION SELECTIVITY AND PERMEABILITY  
 FT SITE 1450 1450 (BY SIMILARITY).  
 FT SITE 1450 1450 CALCIUM ION SELECTIVITY AND PERMEABILITY  
 FT BINDING 1134 1224 (BY SIMILARITY).  
 FT BINDING 1464 1530 TO DIHYDROPYRIDINES (BY SIMILARITY).  
 FT  
 Query Match 19.0%; Score 60; DB 1; Length 2203;  
 Best Local Similarity 34.9%; Pred. No. 65;  
 Matches 22; Conservative 9; Mismatches 18; Indels 14; Gaps 4;  
 QY 1 ENPKHNK--CIQSCNSR---DSYRNA---CHARCNLKVEKCEGEIIPR---P 46  
 DB 843 ENKKNKPEVNOJLANSNKVTIDYGEARDKDPYPCGVPGVEEEDDEDEVEVAPG 902  
 QY 47 RPR 49  
 DB 903 RPR 905  
 RESULT 11  
 ID BARI-CHITE STANDARD: PRT; 174 AA.  
 AC P02849;  
 DT 21-JUL-1986 (rel. 01, Created)  
 DT 21-JUL-1986 (rel. 01, Last sequence update)  
 DT 01-FEB-1994 (rel. 28, Last annotation update)  
 DE BALBIANI RING PROTEIN 1 (GIANT SECRETORY PROTEIN I-A) (GSP-1A)  
 DE (FRAGMENT).  
 GN BRL.  
 OS Chironomus tentans (Midge).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;  
 OC Chironomidae; Chironomidae; Chironominae; Chironomus.  
 RN (1)  
 RP SEQUENCE OF 1-116 FROM N.A. (CLONE PCTBRI-1).  
 RC TISSUE-SALIVARY GLAND;  
 RX MEDLINE=83238361; PubMed=6305953;  
 RA Case S.T., Byers M.R.;  
 RT Repeated nucleotide sequence arrays in Balbiani ring 1 of Chironomus  
 tentans contain internally nonrepeating and subrepeating elements.\*;  
 RL J. Biol. Chem. 258:7793-7799(1983).  
 RN (2)

```

RP SEQUENCE OF 7-174 FROM N.A. (CLONE PCT21).
RC TISSUE-SALIVARY GLAND:
RA Wieslander L., Smeeti J., Daneholt B.;
RT "Evidence for a common ancestor sequence for the Balbiani ring 1 and
RL Balbiani ring 2 genes in Chironomus tentans.";
RC Proc. Natl. Acad. Sci. U.S.A. 79:6956-6960(1982).
CC -1- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
CC STRUCTURE, THE LARVAL TUBE.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: SALIVARY GLAND.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on lists
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-ib.ch).
CC -----
DR EMBL, K00447: AAA28238.1; -.
DR EMBL, J01055: AAA28236.1; -.
DR PIR, A03359; BQICIT.
KW Repeat.
FT NON_TER 1 1
FT DOMAIN 42 85 4 X 11 AA TANDEM REPEATS.
FT REPEAT 42 52 1-1.
FT REPEAT 53 63 1-2.
FT REPEAT 64 74 1-3.
FT REPEAT 75 85 1-4.
FT DOMAIN 124 167 4 X 11 AA TANDEM REPEATS.
FT REPEAT 124 134 2-1.
FT REPEAT 135 145 2-2.
FT REPEAT 146 156 2-3.
FT REPEAT 157 167 2-4.
FT VARIANT 40 40 N -> K (IN CLONE PCT21).
FT VARIANT 57 57 G -> E (IN CLONE PCT21).
FT VARIANT 72 72 K -> R (IN CLONE PCT21).
FT VARIANT 86 86 R -> G (IN CLONE PCT21).
FT NON_TER 174 174
SO SEQUENCE 174 AA; 18920 MW; E86E616C6413D751 CRC64;

Query Match 18.9%; Score 59.5; DB 1; Length 174;
Best Local Similarity 26.5%; Pred. No. 7.4;
Matches 18; Conservative 11; Mismatches 24; Indels 15; Gaps 4;

OY 3 PKHNK--LQSCNERSYSRN--QACHARCNL--KVEKECEGEELPR-----R 47
   1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
DB 82 PRPEKSGAMKRAEKAKARRNGRNMAKCKCTSGKRSKSPKSGKSPREKSKESK 141
   1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:

OY 48 PRPOHPR 55
   111: 1:
DB 142 PRPEKSK 149

RESULT 12
MS11_DROHY STANDARD: PRT; 344 AA.
AC 008695;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE AXONEME-ASSOCIATED PROTEIN MS1101(1).
GN MS1101(1).
OS Drosophila hydei (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyridiacea; Drosophilidae; Drosophila.
RN 11
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE-TESTIS;
RX MEDLINE=94200512; PubMed=8150205;
RA Neesen J., Buemann H., Heinlein U.A.;
```

[illegible]





